



Αναγνώριση Προτύπων

Εργασία Μαθήματος – Εφαρμογή τεχνικών αναγνώρισης ρόλων μηχανικών λογισμικού με χρήση τεχνικών εξόρυξης γνώσης από αποθετήρια ανοιχτού λογισμικού

1. Σκοπός της εργασίας

Σκοπός της εργασίας είναι η ανάπτυξη μιας μεθοδολογίας η οποία εφαρμόζει τεχνικές εξόρυξης γνώσης για την αναγνώριση των διαφορετικών ρόλων που μπορεί να έχει ένας μηχανικός λογισμικού σε έργα ανάπτυξης λογισμικού. Προς την κατεύθυνση αυτή μπορούν να εφαρμοστούν τόσο τεχνικές ταξινόμησης, όσο και ομαδοποίησης.

2. Διαδικασία

Η διαδικασία αντιμετώπισης του προς εξέταση προβλήματος έχει ως εξής:

1. Δείτε τα δεδομένα (dataset) που σας δίνονται και προσπαθήστε να κατανοήσετε τις μετρικές που χαρακτηρίζουν την συνεισφορά του κάθε μηχανικού σε ένα έργο λογισμικού. Στη διαδικασία αυτή θα σας βοηθήσει ιδιαίτερα η περιγραφή των μετρικών, καθώς και οι ενδεικτικές προσεγγίσεις που βρίσκονται στο αρχείο *Problem_Statement_Dataset_2018.pdf*.
2. Επιλέξτε τα ερευνητικά ερωτήματα που σας ενδιαφέρουν και με τα οποία σκοπεύετε να ασχοληθείτε. Στη διαδικασία αυτή μπορείτε να ασχοληθείτε με τα ερευνητικά ερωτήματα που διατυπώθηκαν στην εκφώνηση καθώς και να προσθέσετε δικά σας.
3. Σκεφτείτε προσεκτικά τη μεθοδολογία που θα ακολουθήσετε προσπαθώντας να απαντήσετε στα παρακάτω ερωτήματα:
 - ✓ Ποιός είναι ο στόχος μου;
 - ✓ Γιατί επέλεξα αυτές τις μετρικές;
 - ✓ Η μεθοδολογία που επέλεξα καταλήγει σε ένα ολοκληρωμένο σύστημα ή προϋποθέτει περισσότερα βήματα;
 - ✓ Έκανα κάποια αυθαίρετη παραδοχή και, αν ναι, πού βασίζεται;

4. Enjoy!! 😊

Hint: Για να διαβάσετε τα δεδομένα στην R μπορείτε να χρησιμοποιήσετε την παρακάτω εντολή:

```
dataset = read.csv("dataset.csv", sep=';', header = TRUE, na.strings = c("NA"),  
colClasses = c("character", "character", "numeric", "numeric", "integer", "integer",  
"integer", "numeric", "numeric", "integer", "integer", "integer", "integer",  
"integer", "integer", "numeric", "integer", "integer", "integer", "integer",  
"integer", "integer", "integer", "integer", "integer", "integer", "integer"))
```



Μπορείτε να υλοποιήσετε τη λύση σας σε οποιαδήποτε γλώσσα προγραμματισμού επιλέξετε. Σε περίπτωση επιλογής κάποιας άλλης γλώσσας εκτός της R, προσέξτε ώστε ο τύπος δεδομένων κάθε στήλης να διαβαστεί σωστά (integer, numeric/float, character/string).

3. Ομάδες Εργασίας

Οι ομάδες είναι **2 ατόμων**. Θα τις επιλέξετε στον δικτυακό τόπο του μαθήματος στο e-learning, στην Ενότητα “Εργασίες -> Εργασία Ομάδες”.

Προθεσμία δήλωσης ομάδας: Πέμπτη 10/01/2019, 23:59 (Αυστηρή προθεσμία!)

4. Προθεσμία Υποβολής Εργασίας

Προθεσμία Υποβολής: Κυριακή, 24/02/2019, 23:59

5. Βαθμολογία

Η εργασία προσδίδει **έως και 1,5 μονάδα** επιπλέον στο βαθμό του φοιτητή. Η εργασία θα αξιολογηθεί με βάση τα παρακάτω κριτήρια:

1. Επιλογή των ερευνητικών ερωτημάτων.
2. Τρόπος προ-επεξεργασίας/αρχικής ανάλυσης δεδομένων.
3. Ο αριθμός των πειραμάτων που έχουν διεξαχθεί και οι βελτιστοποιήσεις που εφαρμόστηκαν.
4. Τα αποτελέσματα στα οποία κατέληξε η μεθοδολογία που εφαρμόστηκε.
5. Η πολυπλοκότητα της προσέγγισης που ακολουθήθηκε.
6. Η έγγραφη αναφορά

6. Παραδοτέα

1. Ο **πηγαίος κώδικας** που θα υλοποιήσετε. Η υλοποίηση μπορεί να γίνει σε όποια γλώσσα προγραμματισμού επιλέξετε.
2. Όλα τα **αποτελέσματα** στα οποία καταλήξατε.
3. **Έγγραφο αναφορά**, η οποία θα περιέχει τα εξής:
 - a. Τον τίτλο της εργασίας σας, τα ονόματα και τις ηλεκτρονικές διευθύνσεις των μελών της ομάδας σας.
 - b. Σύντομη περιγραφή του προβλήματος που καλείστε να επιλύσετε.
 - c. Τις λύσεις που δώσατε (παραμέτρους και πειράματα).
 - d. Την αξιολόγηση των αποτελεσμάτων σας.
 - e. Τα συμπεράσματα στα οποία καταλήγετε.
 - f. Προτάσεις για βελτίωση της μεθοδολογίας σας.
4. Τα **τελικά μοντέλα** σας.



ΠΡΟΣΟΧΗ: Η έγγραφη αναφορά σας θα βασίζεται πάνω στο υπόδειγμα (template) που σας έχει δοθεί. Το αρχείο αυτό είναι το **template.doc**.

Το όριο για το σύνολο της εργασίας είναι 5000 λέξεις. Κάθε εικόνα (figure) και κάθε πίνακας προσμετράται για 100 λέξεις και συνυπολογίζεται στο μέγεθος της εργασίας. Οποιαδήποτε υπέρβαση σε μέγεθος θα επηρεάζει αρνητικά τη βαθμολογία.

Για περισσότερες πληροφορίες:

Ανδρέας Λ. Συμεωνίδης

Επ. Καθηγητής

E-mail: asymeon@eng.auth.gr

Παπαμιχαήλ Μιχαήλ

Μεταπτυχιακός Φοιτητής

E-mail: mpapamic@issel.ee.auth.gr